The Political Economy of Heterogeneity and Conflict

Enrico Spolaore and Romain Wacziarg

PWP-CCPR-2017-003

April 6, 2017

California Center for Population Research
On-Line Working Paper Series

The Political Economy of Heterogeneity and Conflict*

Enrico Spolaore
Tufts University and NBER

Romain Wacziarg
UCLA and NBER

March 2017

Abstract

We present a conceptual framework linking cultural heterogeneity to inter-group conflict. When conflict is about control of public goods, more heterogeneous groups are expected to fight more with each other. In contrast, when conflict is about rival goods, more similar groups are more likely to engage in war with each other. We formalize these ideas within an analytical model and discuss recent empirical studies that are consistent with the model's implications.

---

*Spolaore: Department of Economics, Tufts University, Medford, MA 02155-6722, enrico.spolaore@tufts.edu. Wacziarg: UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles CA 90095, wacziarg@ucla.edu. This paper was prepared as keynote lecture for the CESIfo conference on "Disrupted Economic Relationships: Disasters, Sanctions and Dissolutions," organized by Volker Nitsch and Tibor Besedeš and held in Venice, Italy. We thank conference participants for their comments.

1 Introduction

Does heterogeneity of cultural traits and ethnic origins go hand in hand with more conflict between different groups? Social scientists have often held polar views on this question.

At one extreme is the optimistic view that historically heterogeneous populations, by interacting and cooperating with each other, can converge on common norms and values and achieve peaceful and sustainable integration. For instance, versions of this view inspired the functionalist approach to European integration (Haas, 1958 and 1964), as well as broader theories about communication and political cooperation across communities (for example, Deutsch, 1964). Indeed, after the Second World War, Europeans have managed to create common institutions through peaceful integration of a growing and increasingly more diverse set of populations. However, the recent wave of crises and disruptions in Europe - including the Brexit vote in June 2016 and the surge of anti-EU political movements in several countries - has challenged such optimistic assumptions, raising questions about costs and instability associated with political and cultural heterogeneity.[1]

At the other extreme is the pessimistic view that ethnic and cultural dissimilarity prevents cooperation and brings about conflict and wars. This "primordialist" view has a long intellectual pedigree (for instance, Sumner, 1906), but has received renewed attention more recently, especially after the collapse of the Soviet Union. A well-known example of this position is Huntington's (1993 and 1996) "Clash of Civilizations" hypothesis, stressing religious and cultural cleavages as major sources of violent conflict in the post-Cold-War world.

In this chapter, we present a conceptual framework to shed insights on the relation between conflict and cultural heterogeneity. Our central point is that the impact of heterogeneity should depend on whether groups are fighting over the control of public goods or rival goods. Heterogeneous preferences and traits negatively affect the provision of public goods, which are non-rival in consumption and must be shared by all within a jurisdiction, whether one likes them or not. In contrast, diversity across individuals and groups comes with benefits when considering interactions about rival goods, because a diversity of preferences and culture should be associated with lower levels of antagonism over a specific private good. In such cases, it is similarity of preferences that should bring about more conflict.

[1] For recent discussions of the political economy of European integration, stressing the role of heterogeneity costs, see Spolaore (2013, 2015).

This chapter's main idea can be illustrated with a simple example. Consider two people in a room with two sandwiches: a chicken sandwich and a ham sandwich (rival goods). People who share more similar preferences are more likely to want the same kind of sandwich, and possibly to fight over it, while people with more diverse preferences are more likely to be happy with different sandwiches. In contrast, suppose that there is a television set in the room, which both individuals must share (public good). Each can watch television without reducing the other person's utility from watching, but they may disagree over which channel to watch and fight over the remote control. In this case, people with more similar preferences are less likely to fight, because they can agree on the same show.

If this distinction is relevant to understand actual conflict, we should observe more conflict over public goods among groups that are more dissimilar, but more conflict over rival goods among groups that are closer to each other in preferences, values, and cultures. In order to bring these hypotheses to the data, we must be able to measure heterogeneity and distance between different groups and populations. While such measuring is complex and conceptually tricky, there is now a large and growing empirical literature that has made substantial progress on these issues. In our own work on this topic, we have taken a genealogical approach to heterogeneity across different populations (for a recent discussion, see for example Spolaore and Wacziarg, 2016a). The approach is based on the idea that all human populations are related to each other, but some share more recent common ancestors than others, with direct implications for the extent that they are more similar in several relevant characteristics and preferences. As different populations have split from each other over the long time, they have gradually diverged on sets of traits that are transmitted from one generation to the other, including language, values and norms. As a result, on average, populations with a more recent common history have had less time to diverge over such intergenerationally transmitted traits, and tend to be more similar. Henceforth, we can use measures of long-term relatedness between populations to test whether heterogeneity across different groups is associated with more or less conflict among them.

Our own empirical findings on heterogeneity and conflict, focused on international wars (Spolaore and Wacziarg, 2016b), are consistent with the central hypothesis of this chapter. In particular, we have found that, over the past two centuries, sovereign states inhabited by more closely related populations have been more likely to engage in violent conflict over rival goods, such as territories and natural resources (fertile soil in the 19^{th} century, oil in the 20^{th} century).

The rest of this chapter is organized as follows. In Section 2 we provide an analytical framework, capturing our main ideas about the links between intergenerational transmission of preferences, heterogeneity and conflict over rival goods or public goods. In Section 3 we discuss recent empirical studies of conflict that strongly support the implications of our analytical framework. Section 4 concludes.

2 Heterogeneity and Conflict: An Analytical Framework

In this section we present a basic theoretical framework linking intergenerational transmission of preferences, genealogical distance, and the probability of conflict between populations. First, we model the transmission of preferences over time with variation across populations, in order to explain why differences in values, norms and preference are linked to the degree of genealogical relatedness between populations. We show that populations that are more closely related - i.e., at a smaller genealogical distance - tend to have more similar preferences. Second, we model conflict over rival goods, and show that conflict is more likely to arise when different populations care about the same rival goods and resources. Third, we show how the effect of relatedness on conflict changes if the dispute is about control of non-rival goods - i.e. public goods. Finally, we present a generalization of the framework which includes the two previous settings (conflict over rival goods and conflict over non-rival goods) as special cases.

2.1 Intergenerational Transmission of Preferences

Our starting point is a simple model of the intergenerational transmission of preferences over the very long run. Consider three periods: o for origin, p for prehistory, and h for history. In period o there exists only one population: population 0. In period p, the original population splits into two populations: population 1 and population 2. In period h, each of the two populations splits into two separate populations again: population 1 into population 1.1 and population 1.2, and population 2 into population 2.1 and population 2.2, as displayed in Figure 1. In this setting, the genealogical distance $d_g(i, j)$ between population i and population j can be simply measured by the number of periods since they were one population:

$$d_g(1.1, 1.2) = d_g(2.1, 2.2) = 1 \tag{1}$$

and

$$d_g(1.1\ ,\ 2.1) = d_g(1.1\ ,\ 2.2) = d_g(1.2\ ,\ 2.1) = d_g(1.2\ ,\ 2.2) = 2 \tag{2}$$

These numbers have an intuitive interpretation: populations 1.1 and 1.2 are sibling populations, sharing a common parent ancestor (population 1), while populations 2.1 and 2.2 are also sibling populations, sharing a different common parent ancestor (population 2). In contrast, populations 1.1 and 2.1, for example, are cousin populations sharing a common grand-parent ancestor (population 0).

For simplicity, preferences are summarized by two types (A and B). At time o, the ancestral population 0 is either of type A or of type B. For analytical convenience and without loss of generality, we assume that population 0 is of type A with probability 1/2 and of type B with probability 1/2.[2] Populations inherit preferences from their ancestors with variation - a population i' descending from a population i will have preferences of the same type as their parent population i with probability μ, and of the other type with probability $1-\mu$.

We capture the fact that populations inherit preferences from their ancestors by assuming $\mu > 1/2$ and the fact that there is variation (inheritance is not perfect) by assuming $\mu < 1$.[3] Then, on average, populations at a smaller genealogical distance from each other will tend to be more similar in preferences. For instance, the probability that two sibling populations (e.g., 1.1 and 1.2) have identical types is

$$F(\mu) = \mu^2 + (1-\mu)^2 \tag{3}$$

while the probability that two cousin populations (e.g., 1.1 and 2.1) have identical types is

$$G(\mu) = \mu^4 + 6\mu^2(1-\mu)^2 + (1-\mu)^4 \tag{4}$$

It can be easily shown that[4]

$$F(\mu) > G(\mu) \qquad \text{for } \frac{1}{2} < \mu < 1 \tag{5}$$

[2] The qualitative results would not change if we were to assume that the ancestral population is of type A with probability 100% or of type B with probability 100%.

[3] At $\mu = 1/2$, each population would have equal chances of being of either type, independently of the parent population's type, while at $\mu = 1$, each population would be of the same type as their ancestors with 100% probability.

[4] By dividing both $F(\mu)$ and $G(\mu)$ by μ and rearranging terms, the inequality $F(\mu) - G(\mu) > 0$ can be re-written equivalently as

$$2 - 10\mu + 16\mu^2 - 8\mu^3 \equiv f(\mu) > 0$$

It is immediate to verify that the above inequality holds, given that $f(\frac{1}{2}) = f(1) = 0$ and the derivative

$$f'(\mu) = 2(-5 + 16\mu - 12\mu^2)$$

is strictly positive for $1/2 < \mu < 5/6$, zero at $\mu = 5/6$, and negative for $5/6 < \mu \leq 1$.

which implies:

Proposition 1

The probability that two populations are of the same type is decreasing in genealogical distance.

This result plays a key role in our analysis of conflict below.

2.2 Conflict over Rival Goods

Consider two populations (i and j), each forming a sovereign state. For simplicity, we assume that each state is a unified agent, formed by one population with homogeneous preferences.[5]

Suppose that sovereign state i is in control of a valuable prize of type t, from which it obtains the following benefits b_i:

$$b_i = (1 - |t - t_i^*|)R \tag{6}$$

where t_i^* denotes state i's ideal type, and $R > 0$ is the size of the prize. If the prize is of type A, $t = t_A$, and if it is of type B, $t = t_B$. Without loss of generality, we assume that the prize is of type A with probability 1/2 and of type B with probability 1/2. State i's ideal type is also equal to either t_A or t_B. We assume that the state benefits from controlling the prize even if it is not of its favored type, that is

$$|t_A - t_B| < 1 \tag{7}$$

The prize can be interpreted as any valuable good which can be controlled by a sovereign state - e.g. natural resources, land, cities, trade routes, colonies, protectorates, etc. (we return to the interpretation of the model below, when we discuss possible extensions). Sovereign state j also values the prize, and would gain positive benefits if it could control the prize. State j's benefits b_j from controlling the prize are

$$b_j = (1 - |t - t_j^*|)R \tag{8}$$

State j can try to obtain control over the prize by challenging state i - that is, state j can take two actions: "challenge" state i (C) or "not challenge" (NC). If state j chooses action NC, state i keeps full control over the prize, and obtains a net utility equal to b_i, while state j obtains net

[5] That is, we abstract from the possibility that states may include mixed populations with different preferences. However, in the empirical analysis reviewed in Section 3, we take into account population heterogeneity within states when computing distance between states.

benefits equal to zero. If state j challenges state i for the possession of the prize, state i can respond either with "fight" (F) or "not fight" (NF). If state i does not fight, state j obtains control of the prize, and net benefits equal to b_j, while state i obtains net benefits equal to zero.

If state i decides to fight in response to the challenge, a war takes place.[6] When a war occurs - i.e., when actions $\{C, F\}$ are taken, the probability that state i wins, denoted by π_i, is a function of the two states' relative military capabilities (denoted respectively by M_i and M_j):

$$\pi_i = \frac{M_i}{M_i + M_j} \tag{9}$$

while the probability that state j wins the war is obviously $1 - \pi_i$. This is an instance of "ratio" contest success function. In general, the literature on the technology of conflict assumes that the probability of success is a function of either the ratio or the difference between military capabilities (for a general discussion, see Garfinkel and Skaperdas, 2007).[7]

Ex ante, each state obtains an expected utility respectively given by

$$U_i = \pi_i b_i - c_i \tag{10}$$

$$U_j = (1 - \pi_i) b_j - c_j \tag{11}$$

where $c_i > 0$ and $c_j > 0$ denote the respective costs of going to war. The extensive form of the game is illustrated in Figure 2.

It is immediate to show that:

Lemma

War is a sub-game perfect equilibrium if and only if $\min\{U_i, U_j\} \geq 0$. *War is the unique sub-game perfect equilibrium when* $\min\{U_i, U_j\} > 0$.[8]

[6] In the Appendix, we present an extension in which peaceful bargaining is possible, in alternative to war, when state j challenges and state i responds to the challenge.

[7] The choice of specification in this paper is inconsequential because we treat military capabilities as exogenous. A straighforward extension would be to endogenize military capabilities. The extension could strengthen the link between relatedness and probability of conflict, insofar as states with similar preferences might face more similar incentives to invest in military capabilities, all other things being equal.

[8] When $U_i > 0$ and $U_j = 0$, two sub-game perfect equilibria exist: $\{C, F\}$ and $\{NC, F\}$. When $U_i = 0$ and $U_j > 0$, there are also two sub-game perfect equilibria: $\{C, F\}$ and $\{C, NF\}$. When $U_i = U_j = 0$, three equilibria may occur: $\{C, F\}$, $\{C, NF\}$ and $\{NC, F\}$. When $\min\{U_i, U_j\} < 0$ the only sub-game perfect equilibria are peaceful. If $U_i < 0$, the only sub–game perfect equilibrium is $\{C, NF\}$. If $U_i > 0$ and $U_j < 0$, the only sub-game perfect equilibrium is $\{NC, F\}$. Finally, when $U_i = 0$, and $U_j < 0$ there are two (peaceful) equilibria: $\{NC, F\}$ and $\{C, NF\}$.

Proof: When $U_i > 0$ and $U_j = 0$, two sub-game perfect equilibria exist: $\{C, F\}$ and $\{NC, F\}$. When $U_i = 0$ and $U_j > 0$, there are also two sub-game perfect equilibria: $\{C, F\}$ and $\{C, NF\}$. When $U_i = U_j = 0$, three equilibria may occur: $\{C, F\}$, $\{C, NF\}$ and $\{NC, F\}$. When $\min\{U_i, U_j\} < 0$ the only sub-game perfect equilibria are peaceful. If $U_i < 0$, the only sub–game perfect equilibrium is $\{C, NF\}$. If $U_i > 0$ and $U_j < 0$, the only sub-game perfect equilibrium is $\{NC, F\}$. Finally, when $U_i = 0$, and $U_j < 0$ there are two (peaceful) equilibria: $\{NC, F\}$ and $\{C, NF\}$. QED.

We are now ready to investigate how similarity in preferences between the two states affect the probability of war. To simplify the analysis, we assume equal capabilities ($M_i = M_j = M$) and costs ($c_i = c_j = c$). Let $P(i, j)$ denote the probability of a war between state i and state j.

A war *never* occurs ($P(i, j) = 0$) if each state's expected utility from going to war is negative even when the prize is of its preferred type. This happens at a very high cost of war:

$$c > \frac{1}{2}R \tag{12}$$

In contrast, a war *always* occurs ($P(i, j) = 1$) if each state's expected utility from going to war is positive even when the resource is not of its favored type. This happens at a very low cost of war:

$$c < \frac{1}{2}R(1 - |t_A - t_B|) \tag{13}$$

Therefore, we focus on the more interesting case when war may occur with probability between 0 and 1 (that is, $0 < P(i, j) < 1$), which happens when the cost of war takes on an intermediate value:[9]

$$\frac{1}{2}R(1 - |t_A - t_B|) < c < \frac{1}{2}R \tag{14}$$

Under these assumptions, a war will occur if and only if the two states have the same preferred type, and that type is equal to the type of the prize under dispute - that is, $t_i^* = t_j^* = t$. If the two states had always identical preferences, the probability of a war would be 1/2. This would occur, for instance, if preferences were transmitted without variation across generations: $\mu = 1$. In contrast, if the preferences of each state were independently distributed, with each state having a 50% chance of preferring type A to type B (and vice versa), the probability of war would be 1/4. This would occur, for instance, if preferences were transmitted purely randomly across generations: $\mu = 1/2$.

[9] To simplify the analysis, we do not consider the knife-edge cases $c = \frac{1}{2}R$ and $c = \frac{1}{2}R(1 - |t_A - t_B|)$, when it's possible that $\min\{U_i, U_j\} = 0$, implying that one or both states may be indifferent between war and peace, and multiple equilibria may trivially occur, as detailed in the proof to the Lemma above.

In general, for $1/2 < \mu < 1$, the expected probability of war between states i and j would depend on the degree of relatedness (genealogical distance) of their populations. For two states i and j with $d_g(i,j) = 1$ - i.e., states formed by sibling populations - the probability that both states' type is the same as the prize under dispute is half the probability that both states have the same preferences, that is

$$P\{i, j \mid d_g(i,j) = 1\} = \frac{F(\mu)}{2} = \frac{\mu^2 + (1-\mu)^2}{2} \tag{15}$$

By the same token, for states such that $d_g(i,j) = 2$ - i.e., states formed by cousin populations -, the probability that both states' type is equal to the type of the prize is

$$P\{i, j \mid d_g(i,j) = 2\} = \frac{G(\mu)}{2} = \frac{\mu^4 + 6\mu^2(1-\mu)^2 + (1-\mu)^4}{2} \tag{16}$$

As already shown in the previous section, $F(\mu) > G(\mu)$ for all $1/2 < \mu < 1$. Therefore, it immediately follows that

$$P\{i, j | d_g(i,j) = 1\} > P\{i, j \mid d_g(i,j) = 2\} \tag{17}$$

which we can summarize as our main result:

Proposition 2

States with more closely related populations (smaller genealogical distance) are more likely to go to war with each other.

An illustration of the model can be provided with a simple spatial example. Assume that space is unidimensional. Three states divide the territory among themselves as in Figure 3, with the border between state i and state j at point x, and the border between state j and state i' at point y. Assume that state i and state j are of type A, and state i' is of type B. The parameters are such that equation (14) is satisfied. Now, consider the territory between x' and x. If that territory is of type B, state j will not challenge state i for its possession, but if that territory is of type A, a war will occur. In contrast, consider the territory between y and y'. If that territory is of type B, state j will not challenge state i' for its possession, while if it is of type A, state j will challenge state i', and state i' will surrender it peacefully. In either case, no conflict will occur.

This example illustrates how the probability of conflict between states in similar geographical settings varies because of preferences over the prize: states with more similar preferences are more likely to go to war with each other, other things being equal. In this example, the prize is a contiguous territory, but similar effects would hold for control over non-contiguous territories (colonies, protectorates, ports and harbors along trade routes), or over other rival issues about which states may care with different intensities (for instance, monopoly rights over trade, fishing or other valuable sources of income in specific waters or regions). History is abundant with examples of populations that fought over specific rival goods (territories, cities, religious sites) because they shared a common history and common preferences, inherited with variation from their ancestors. For instance, genetically close populations (Jews and Arabs) who share similar preferences over Jerusalem have fought and continue to fight over the control of that rival good. In general, we can expect that populations may share more similar preferences over specific types of land and resources because they have inherited similar tastes and demand functions (as in the example about Jerusalem), or because they have inherited similar technologies and methods of production, or both.[10]

2.3 Conflict over Public Goods

In our basic model the prize is a rival and excludable good: either one or the other state obtains full control, and the population in the state without control receives no net benefit. How would our results change if the prize were a public good (non-rival and non-excludable in consumption)? Then, state j would obtain some external benefits when state i is in control of the good, and vice versa. In itself, this extension would only reduce the likelihood of war, because the externalities would reduce the gap in utility between controlling and not controlling the good. However, the implications would change dramatically if we also allowed the state in control to select the characteristics or type of the public good.

In our basic model the prize is a rival and excludable good, and the type of the rival good is given (that is, it cannot be changed by either player). We now consider the polar case, where the prize is a pure public good, non-rival in consumption, and the player in control can choose whether the public good is of type A or B. In this case, we refer to "players" rather than "states," consistently with our view that conflict over types of public goods is more likely to occur among agents engaged in intrastate conflict rather than interstate conflict.

Now, the conflict is not about controlling access to the good (both players benefit from the

[10] In principle, similarities in technology could affect the probability of conflict not only by affecting preferences over rival goods, but also, more directly, by affecting military capabilities (more similar populations may be more similar in military technologies and hence capabilities, other things being equal).

good no matter who "owns" it), but about controlling the type (e.g., the characteristics of a public policy or service): the winner will select his/her favored type of public good. Hence, utilities from the public good are given as follows:

a) If player i and player j are of the same type, both obtain maximum benefits R from the good no matter who is in control:

$$b_i = b_j = R \tag{18}$$

b) If the two players are of different types, and player i is in control of the public good, the respective benefits are

$$b_i = R \tag{19}$$

$$b_j = (1 - |t_A - t_B|)R \tag{20}$$

c) Conversely, if the two players are of different types, and player j is in control of the public good, we have:

$$b_i = (1 - |t_A - t_B|)R \tag{21}$$

$$b_j = R \tag{22}$$

Now, there is no reason for conflict between two players of the same type: if player i is of the same type as player j, player j will obtain the same utility as if he/she were in control of the good. In contrast, if player i is of a different type, player j could increase his/her utility by seizing control of the good and changing the type. Hence, a necessary condition for war is that the players are of *different* types. Then, player i's expected utility from going to war is

$$\pi_i R + (1 - \pi_i)[1 - |t_A - t_B|]R - c_i \tag{23}$$

and he/she will prefer to fight when

$$\pi_i R + (1 - \pi_i)(1 - |t_A - t_B|)R - c_i > (1 - |t_A - t_B|)R \tag{24}$$

which can be re-written as

$$\pi_i(|t_A - t_B|)R - c_i > 0 \tag{25}$$

By the same token, player j's will prefer war over not challenging the other player for control when

$$(1 - \pi_i)(|t_A - t_B|)R - c_j > 0 \tag{26}$$

In the symmetric case ($\pi_i = 1/2$ and $c_i = c_j = c$) the two conditions become

$$c < (|t_A - t_B|)\frac{R}{2} \tag{27}$$

If the condition above is satisfied - i.e., the war costs are small enough - the probability that player i and player j engage in conflict is equal to the probability that they are not of the same type. For sibling populations ($d_g(i,j) = 1$), the probability that they are not of the same type is:

$$P\{i, j \mid d_g(i,j) = 1\} = 1 - F(\mu) = 1 - [\mu^2 + (1-\mu)^2] \tag{28}$$

while for cousin populations ($d_g(i,j) = 2$), the probability that both are of different types is

$$P\{i, j \mid d_g(i,j) = 2\} = 1 - G(\mu) = 1 - [\mu^4 + 6\mu^2(1-\mu)^2 + (1-\mu)^4] \tag{29}$$

As we have shown above, $F(\mu) > G(\mu)$ for all $1/2 < \mu < 1$,which immediately implies

$$P\{i, j|d_g(i,j) = 1\} < P\{i, j \mid d_g(i,j) = 2\} \tag{30}$$

that is, we have:

Proposition 3

When conflict is about the control of public-good types, the probability of violent conflict is higher between groups that are less closely related.

The intuition is straightforward: Suppose conflict is not about control of the public good *per se*, but about determination of its type. Then, more closely related populations, sharing more similar preferences about the characteristics of the public good, are less likely to engage in conflict. In contrast, populations that are historically and culturally more distant tend to disagree more over the type of public good.

2.4 A General Framework

The two basic models - conflict over pure rival goods and conflict over pure public goods - can be viewed as two special cases of a more general framework where: (a) there may be externalities in consumption, and (b) the player in control of the prize may be able to change the good's type. Formally:

(a) when player i is in control of the prize, player j's benefits are δb_j and when player j is in control, player i's benefits are δb_i where $0 \leq \delta \leq 1$.

(b) when a player is in control of the prize of type A, he/she can change the type to B (and vice versa, can change type B to type A), with probability γ ($0 \leq \gamma \leq 1$).

Our basic model of conflict over rival goods is the case $\delta = \gamma = 0$, while the model of conflict over public goods is the polar case $\delta = \gamma = 1$.

In general, two players with the same preferences will go to war with each other at low levels of δ (for all γ), while two players with different preferences will go to war with each other for high levels of γ, when $\delta > 0$. These results generalize the insights from the basic models: similarity in preferences leads to more conflict over goods with zero or low externalities (low δ), while dissimilarity in preferences leads to more conflict when agents in control of a non-rival good ($\delta > 0$) can change the good's type (high γ). Formally, we have:

Proposition 4

For all γ, there exists a critical $\delta^* = 1 - \dfrac{2c}{R}$ such that two players of the same type will go to war for $\delta < \delta^*$ and will not go to war for $\delta > \delta^*$.[11]

Proof. Two players of the same type $X (X = A, B)$ will not go to war over a good of type X if

$$\frac{1}{2}R + \frac{1}{2}\delta R - c < \delta R \tag{31}$$

and will not go to war over a good of type $Y \neq X$ if

$$\frac{1}{2}[\gamma R+(1-\gamma)(1-|t_A-t_B|)R]+\frac{1}{2}\delta[\gamma R+(1-\gamma)(1-|t_A-t_B|)R]-c < \delta[\gamma R+(1-\gamma)(1-|t_A-t_B|)R] \tag{32}$$

which can be re-written, respectively, as:[12]

$$\delta < 1 - \frac{2c}{R} \tag{33}$$

and

$$\delta < 1 - \frac{2c}{R[1-(1-\gamma)|t_A - t_B|]} \tag{34}$$

For all $0 \leq \gamma \leq 1$, we have

$$1 - \frac{2c}{R} \geq 1 - \frac{2c}{R[1-(1-\gamma)|t_A - t_B|]} \tag{35}$$

[11] Multiple equilibria, with and without conflict, exist in the knife-edge case $\delta = \delta^*$.

[12] For $\delta = 0$, the condition below reduces to the condition for war in the case of pure rival goods: $c < \frac{1}{2}R$.

Therefore, for all $\delta > \delta^* \equiv 1 - \frac{2c}{R}$ we also have $\delta > 1 - \frac{2c}{R[1-(1-\gamma)|t_A - t_B|]}$, and no war ever takes place between two players with the same preferences. In contrast, for $\delta < \delta^*$, the two players will go to war. QED.

In contrast, conflict between players with different preferences is characterized by the following proposition:

Proposition 5

For all $\delta > 0$, two players with different preferences will go to war for $\gamma > \gamma^*$, where[13]

$$\gamma^* = \frac{1}{|t_A - t_B|} \min\{1 - \frac{1}{\delta}(1 - \frac{2c}{R}) \; ; \frac{2c}{R} - (1-\delta)[1 - |t_A - t_B|]\} \tag{36}$$

Proof. When two players have different preferences, the player whose preferred type is the same as the prize will go to war if

$$\frac{1}{2}R + \frac{1}{2}\delta[\gamma(1 - |t_A - t_B|)R + (1-\gamma)R] - c > \delta[\gamma(1 - |t_A - t_B|)R + (1-\gamma)R] \tag{37}$$

while the other player will go to war if

$$\frac{1}{2}[\gamma R + (1-\gamma)(1 - |t_A - t_B|)R] + \frac{1}{2}\delta(1 - |t_A - t_B|)R - c > \delta(1 - |t_A - t_B|)R \tag{38}$$

The above equations can be re-written as

$$\gamma > \frac{1 - \frac{1}{\delta}(1 - \frac{2c}{R})}{|t_A - t_B|} \tag{39}$$

and

$$\gamma > \frac{\frac{2c}{R} - (1-\delta)(1 - |t_A - t_B|)}{|t_A - t_B|} \tag{40}$$

Both conditions hold if $\gamma > \gamma^*$. QED.

3 Empirical Evidence on Heterogeneity and Conflict

Our conceptual framework implies that conflict over rival goods is likely to be more severe among groups that are more similar in terms of culture, preferences and ethnic origins, while the opposite should occur when conflict is about public goods. What do the data say?

[13] In the case $\delta = \gamma = 1$, the condition below reduces to the condition for war in the case of pure public goods: $c < (|t_A - t_B|)\frac{R}{2}$

3.1 Evidence on Heterogeneity and Civil Conflict

Conflicts over public goods and government characteristics are more likely to emerge among groups that belong to the same political jurisdiction and therefore share non-rival and non-excludable goods and policies by institutional design. Consequently, we can expect that conflict over public goods and policies should play an important role in many (but not all) civil conflicts.

This observation is consistent with empirical work associating ethnic polarization (a measure that captures distance between groups within a country) with conflict over public goods. Of particular note is the empirical study of ethnicity and intrastate conflict by Esteban, Mayoral and Ray (2012), building on theoretical work by Esteban and Ray (2011). In their theoretical framework, Esteban and Ray (2011) and Esteban, Mayoral and Ray (2012) also draw a distinction between public goods and private goods. In their model, a central role is played by three indices, measuring polarization, fractionalization and cohesion. The weight of these indices in explaining conflict intensity depends on the particular nature of each conflict. When group cohesion is high, ethnic polarization increases conflict if the prize is public and fractionalization increases conflict if the prize is private. In their empirical analysis, they use measures of ethnolinguistic polarization based on linguistic distances between groups, building on Fearon (2003), and find that linguistic polarization increases civil conflict over public goods. Such measures can be interpreted in terms of our theoretical model, because linguistic trees capture long-term relations between populations and are correlated with measures of historical and cultural relatedness (see Spolaore and Wacziarg, 2016a). We will return to these measures of linguistic distance and their connection with other measures of cultural distance when discussing our own empirical work below. Overall, the effects of linguistic distance and polarization found by Esteban, Mayoral and Ray (2012) are entirely consistent with the implications of our basic hypothesis that less closely related groups are more likely to fight over the control of public goods.

Desmet, Ortuño-Ortín and Wacziarg (2012) also find a significant impact of linguistic diversity on civil conflict. In their empirical estimates more linguistic diversity within a country is associated with more civil conflict and worse political economy outcomes, such as the provision of public goods, governance, and redistribution. Their analysis is focused on detecting the effects of heterogeneity at different levels of linguistic aggregation. Interestingly, they find that deep cleavages, originating thousands of years ago, are the better predictors of conflict across linguistically heterogeneous groups that share the same country. These findings strongly support the central hypotheses of this chapter. In Desmet, Ortuño-Ortín and Wacziarg (2016), the same authors show that civil conflict is more likely when ethnic divisions are reinforced by cultural cleavages, i.e. differences in preferences, values and norms as revealed in the World Value Survey. This evidence too is consistent with our model, where differences in preferences drive the positive relationship between genealogical relatedness and the likelihood of civil conflict.

Long-term measures of diversity within populations, based on genetic data, are at the center of the study by Arbatli, Ashraf and Galor (2015). Using genetic diversity within each country, they find that more diverse populations are more likely to engage in civil conflict. These results are consistent with our theoretical framework, insofar as civil conflict among people with more diverse preferences is about public goods and policies rather than rival goods.

In sum, recent studies have found significant empirical evidence linking long-term measures of diversity to civil conflict, especially over public goods and policies. This evidence strongly support the hypothesis illustrated in our conceptual framework.

It must be noted, however, that evidence of a positive relationship between heterogeneity and civil conflict does not imply that, in general, cultural and ethnic distance should always be associated with a lower probability of civil conflict, independently of what the different groups within a country are fighting about. In principle, groups engaged in civil and ethnic conflict may also fight over rival goods. Then, according to our conceptual framework, more similar groups would be expected to fight more with each other, and heterogeneity of traits and preferences could in principle have a pacifying effect. In general, our framework predicts that, insofar as civil conflicts are about a complex mix of disputes over rival and public goods, one should expect *ambiguous* effects of heterogeneity on civil conflict, depending on the extent that specific civil conflicts are about rival goods or non-rival goods (public goods). This theoretical ambiguity can shed some light on the ongoing debate on the role of ethnic divisions in causing conflict within countries - e.g., see the path breaking contributions by Fearon and Laitin (2003) and Montalvo and Reynal-Querol (2005).

That said, the more recent evidence on the determinants of civil conflict provided by Esteban, Mayoral and Ray (2012), Desmet, Ortuño-Ortín and Wacziarg (2012), and Arbatli, Ashraf and Galor (2015) suggest two observations. First, when civil conflict is unequivocally about public goods, more heterogeneity is unambiguously associated with more conflict, as predicted by our analytical framework. Secondly, more heterogeneity - measured for instance by greater ethnolinguistic distance - seems to be empirically associated with more civil conflict and worse political economy outcomes within each country. In other words, in the observed historical record, the net impact of long-term cultural heterogeneity on the propensity for civil conflict is positive: more heterogeneity, more conflict. In light of our framework, we can interpret this empirical regularity as consistent with a large role for conflict over public goods and policies. In other words, when heterogenous groups fight with each other *within a country*, chances are that they are disagreeing about the fundamental traits and characteristics of their common government and common policies. More research is necessary, however, to detect the extent that different civil conflicts happen to be about public goods or rival goods, and the consequences of this distinction when estimating the impact of different measures of heterogeneity on civil conflict.

3.2 Evidence on Heterogeneity and International Conflict

If more heterogeneity is typically (but not always) associated with more conflict in the case of civil conflict, what do the data say about international conflict? In principle, international conflict could involve both rival and non-rival goods. However, in contrast to the case of civil conflict, the importance of a "public-goods effect" is likely to be much lower or absent when sovereign states fight with each other. Even though disagreements about the provision of public goods and policies may also emerge among different governments - i.e., how to fight international terrorist threats, global climate change, or financial instability - historically interstate militarized conflicts have been mostly about control of rival and excludable goods, such as territories, cities, and natural resources. The view that international conflict is closely linked to disputes over territories and resources is emphasized, for instance, by Caselli, Morelli and Rohner (2014), who cite the results in Tir et al. (1998) and Tir (2003) that 27% of all territorial changes between 1816 and 1996 involved full-blown military conflict, and 47% of territorial transfers involved some level of violence. Caselli, Morelli and Rohner (2014) also cite Weede's (1973, p. 87) statement that "the history of war and peace is largely identical with the history of territorial changes as results of war."

In our empirical work on international conflict we found that the evidence is unambiguously supportive of the hypothesis that culturally more similar populations fight more with each other - that is, the *opposite* of what would be implied by a "Clash of Civilizations" hypothesis. In what follows we discuss some of the empirical findings reported in Spolaore and Wacziarg (2016b), in light of the conceptual framework presented in this chapter.

3.2.1 Measures of Cultural Distance between Populations

How can we measure distance in cultural traits and preferences between societies? In our empirical analysis, we use a genealogical approach to heterogeneity, consistent with the insights of the model presented above. Specifically, we use genetic distance, which captures the length of time since two populations became separated from each other. The basic idea behind the use of genetic distance as a way to measure cultural distance between populations is that human traits - not only biological but also cultural - are mostly transmitted from one generation to the next, with variation. Therefore, the longer two populations have drifted apart, the greater the differences in cultural traits and preferences between them.

Genetic distance is not the only measure that captures distance in intergenerationally-transmitted traits. A closely related measure is linguistic distance, which is also based on a trait that is mostly transmitted from one generation to the next over time, even though individuals and entire populations have sometime changed their language due to conquest or other factors. Another cultural trait that is mostly transmitted intergenerationally is religion - although in this case also, people can change their religious beliefs. In sum, linguistic and religious distance provide alternative measures of differences in cultural traits that are transmitted with variation from one generation to the next. Another class of distance between populations can be constructed directly by measuring specific differences in cultural traits, values, norms, and attitudes, as revealed by surveys, such as the World Values Survey (WVS). All these different traits are in large part transmitted intergenerationally over time, so we should expect that the various classes of measure based on these traits (genetic distance, linguistic distance, religious distance, cultural distance based on surveys), while distinct from each other, to be positively correlated. This is indeed what we find in Spolaore and Wacziarg (2016a), where we further elaborate on the complex links between various measures of historical and cultural distance between populations, and analyze the empirical relationships between them.

In our empirical study of international conflict (Spolaore and Wacziarg, 2016b), we used three measures of genealogical/cultural distance between countries, based respectively on genetic distance, linguistic distance, and religious distance, to analyze the determinants of interstate conflict. Below, we briefly describe the construction the three measures. More detailed discussions can be found in Spolaore and Wacziarg (2009, 2013, and 2016a).

The data on genetic distance comes from Cavalli-Sforza et al. (1994). The set of world populations from that dataset is matched to ethnic groups from Alesina et al. (2003). In order to account for the fact that modern countries include groups with different ancestry and ethnic origins, we constructed a measure of weighted genetic distance.[14] Assuming that country i is composed of populations $m = 1...M$ and country j is composed of populations $n = 1...N$, and denoting by s_{1m} the share of population m in country i (similarly for country j) and d_{mn} the distance between populations m and n, the weighted F_{ST} genetic distance (GD_{ij}) between countries i and j is defined as:

$$GD_{ij} = \sum_{m=1}^{M} \sum_{n=1}^{N} (s_{im} \times s_{jn} \times gd_{mn}) \tag{41}$$

where s_{km} is the share of group m in country k, gd_{mn} is the F_{ST} genetic distance between groups m and n. This measure represents the expected genetic distance between two randomly selected individuals, one from each country.

To address concerns that current genetic distance may be endogenous with respect to past wars, as well as possible bias resulting from errors in matching populations to countries for the current period, we also matched countries to their populations in the year 1500, before the great migrations following European explorations and conquests. For instance, for 1500 Australia is matched to the Australian Aborigines rather than to the English. We employ such measure of genetic distance based on the 1500 match as an instrument for current genetic distance.

To measure linguistic distance between countries, following Fearon (2003), we use linguistic trees from Ethnologue. We compute the number of common linguistic nodes between languages in the world, a measure of their linguistic similarity. The linguistic tree in this dataset involves up to 15 nested classifications, so two countries with populations speaking the same language will share 15 common nodes.[15] Using data on the distribution of each linguistic group within and across countries, from the same source, we compute a measure of the number of common nodes shared by languages spoken by plurality groups within each country in a pair. Again, to take into account the presence of groups speaking different languages within a country, we computed a weighted measure

[14] We also constructed the distance between the plurality ethnic groups of each country in a pair - that is, the groups with the largest shares of each country's population. Genetic distance based on plurality groups is highly correlated with weighted genetic distance (the correlation is 93.2%). In our empirical analysis we prefer to use weighted genetic distance because it is a more precise measure of average genetic distance between countries.

[15] We have also used a separate measure of linguistic distance, based on lexicostatistics, from Dyen, Kruskal and Black (1992). This is a more continuous measure than the one based on common nodes, but it is only available for countries speaking Indo-European languages. Using the weighted measure of cognate distance led to effects very similar to those obtained when controlling for the Fearon measure, albeit on a much smaller sample of countries.

of linguistic similarity, representing the expected number of common linguistic nodes between two randomly chosen individuals, one from each country in a pair, analogously to the above formula for weighted genetic distance in equation (41). Finally, these measures of linguistic similarity are transformed in an index of linguistic distance (LD_{ij}):

$$LD_{ij} = \sqrt{\frac{(15 - \text{\# Common Linguistic Nodes})}{15}} \tag{42}$$

To measure religious distance we use a family tree of world religions, analogous to the family tree of languages used to compute linguistic distance. The religious nomenclature is obtained from Mecham, Fearon and Laitin (2006). In the tree, we start with three separate branches, one for the monotheistic religions of Middle Eastern origin, one for Asian religions, and a third for a residual category. Then, each branch is subdivided into finer groups - e.g., Christians, Muslims, Jews for the first group, and so on. The number of common classifications - up to 5 in this dataset - captures religious similarity. We match religions to countries using Mecham, Fearon and Laitin's (2006) data on the prevalence of religions by country. The data about common religious nodes are transformed, analogously to what we did for linguistic distance in equation (42). Therefore, we obtain a measure of religious distance (RD_{ij}) between countries:

$$RD_{ij} = \sqrt{\frac{(5 - \text{\# Common Religious Nodes})}{5}} \tag{43}$$

Correlations between measures of genetic, linguistic and religious distances are positive, but not very large, as each reflects a different set of traits that are transmitted intergenerationally with variation. In our dataset of country pairs, the correlation between weighted genetic distance GD_{ij} bears a correlation of 0.201 with weighted linguistic distance LD_{ij} and 0.172 with weighted religious distance RD_{ij}, while the correlation between weighted LD_{ij} and weighted RD_{ij} is 0.449.

3.2.2 Empirical Results

The data on interstate conflict is a 1816-2001 panel from the Correlates of War Project (Jones et. al., 1996 and Faten et al. 2004). In any given year, the indicator of conflict takes on a value from 0 for no militarized conflict to 5 for an interstate war involving more than $1,000$ total battle deaths. As in several other contributions in the literature on interstate conflict, we define a dummy variable equal to 1 if the intensity of militarized conflict is equal to or greater than 3, zero otherwise. In our cross-sectional analysis, we look for pairs that were ever involved in a conflict over the time period 1816-2001.

Our baseline cross-sectional regression specification is:

$$C_{ij} = \beta_1 X_{ij} + \beta_2 GD_{ij} + \eta_{ij} \tag{44}$$

where the vector X_{ij} contains controls such as a contiguity dummy, measures of geodesic distance, longitudinal and latitudinal distance, several other indicators of geographic isolation, and dummy variables indicating whether the countries in a pair were ever part of the same polity and were ever in a colonial relationship. The equation is estimated using probit, clustering standard errors at the country-pair level. Marginal effects are evaluated at the mean of the independent variables. In addition to these marginal effects we also report the standardized magnitude of the effect of genetic distance: the effect of a one standard deviation change in genetic distance as a percentage of the mean probability of conflict. To improve readability, the coefficients are multiplied by 100 in all tables.

Table 1 presents estimates of the coefficients in equation (44) using various specifications. Column (1), the univariate regression, shows a strong negative relationship between weighted genetic distance and the incidence of international conflict. In terms of magnitude, a one standard deviation change in genetic distance (0.068) is associated with a 68.81% decline in the percentage probability of conflict between 1816 and 2001. Needless to say, this estimate is likely to be tainted by omitted variables bias.

In column (2) we introduce eight geographic controls (capturing potential geographic barriers to militarized conflict) and two measures of colonial past. The estimated effects of these measures usually have the expected signs (more distance, less conflict). While the effect of genetic distance is reduced by the inclusion of these controls, it remains negative and highly significant both statistically and economically: a standard deviation increase in genetic distance reduces the probability of conflict by 23.84% relative to the mean.

Column (3) addressed endogeneity and measurement error by instrumenting for modern genealogical distance using genetic distance between populations as they were in the year 1500. Matching countries to genetic groups is much more straightforward for 1500, while genetic distance in 1500 is unlikely to be causally affected by conflicts between 1816 and 2001. The IV results are even stronger: the standardized effect of genealogical distance rises to 36.79%, compared to the estimate of column (2). The effect of genetic distance also remains significant in column (4), where we limit the analysis to countries that are not geographically contiguous.

Finally columns (5) and (6) show the determinants of full-blown wars. That is, here the dependent variable is equal to one if and only if the pair ever experienced a conflict of intensity equal to 5, corresponding to violent conflicts with more than $1,000$ total battle deaths, over the sample period.[16] As before we find that a greater genetic distance has a pacifying effect: a standard deviation increase in genetic distance reduces the probability of ever having experienced a war by 20.57% of this variable's mean. As shown in column (6), the standardized magnitude of the effect rises when we instrument using genetic distance in the year 1500. In Spolaore and Wacziarg (2016b), we further explored the robustness of these results to the inclusion of additional geographic controls in the regression, finding that the baseline results discussed here were not affected.

Table 2 includes the effects of linguistic distance and religious distance. We start in column (1) with the baseline estimates using the new sample for which all variables are available (we lose about 24% of the sample de to unavailable data on linguistic and religious distances). These baseline estimates are similar to those reported in Table 1. When adding linguistic distance and religious distance, the coefficient on genetic distance does not change much. Linguistic distance is not significant when controlling for genetic distance, while religious distance has a negative and significant effect on conflict. The effect of religious distance is consistent with our hypothesis that more similar populations are more likely to fight with each other Religion is an important trait that is transmitted intergenerationally and makes populations more or less related to each other. Populations that share more similar religions are more likely to care about the same holy sites and territories (e.g., Jerusalem) and therefore more likely to fight with each other.

Table 3 presents direct evidence in support of the hypothesis that more similar countries are more likely to fight over rival goods. In column (2) we document a negative interaction between genetic distance and a dummy for oil (1 if at least one of the country has oil, zero otherwise), showing that more similar countries were more likely to fight over oil between 1945 and 2001.[17] Analogous effects are documented in columns (4) and (6) for temperate climate and fertile soil. Those effects hold for conflicts that took place between 1816 and 1900, when agriculture still played a more central role in the world economy.

In sum, the evidence on the determinants of international conflict is strongly supportive of a central role for conflict over rival goods between culturally more similar populations. The evidence

[16] Only 2.1% of the country pairs in our sample ever experienced a war, as defined here, between 1816 and 2001.

[17] Details of the empirical strategy are explained in Spolaore and Wacziarg (2016b).

is also suggestive of an interaction between the two effects highlighted in our conceptual framework. In addition to a direct effect stemming from conflict over rival goods, international conflict (or lack of international conflict) can be influenced by the fact that rulers anticipate the heterogeneity costs associated with conquering populations that are dissimilar from those they already rule. In other words, rulers may care about "winning the peace" after winning the war, and therefore they may be more likely to fight over territories inhabited by people with whom it would be easier to share common public goods and policies after the war. In contrast, rulers may be more willing to allow more heterogeneous populations to become independent without violent conflict. In fact, as shown in Spolaore and Wacziarg (2016b), historically the process of decolonization and independence was more likely to take place peacefully rather than violently when it involved populations that were culturally more distant from the colonial power. Such evidence is consistent with the view that rulers are more likely to fight over a territory when it is inhabited by populations more similar to their own subjects, because more heterogeneous populations involve higher costs to provide common public goods and policies.

4 Concluding Remarks

At the beginning of this chapter we mentioned two polar views regarding the relationship between heterogeneity and conflict: an optimist view and a pessimist (or "primordialist") view. Our theoretical and empirical analysis implies that neither polar view is correct. The optimist view underplays the risks of conflict and disruption when heterogenous populations must share common public goods and policies within a given jurisdiction. The evidence strongly suggest that deeply-rooted cultural divergence can lead to civil conflict over public goods and government characteristics.

However, there is no reason to believe that heterogeneous populations are always bound to fight with each other. In fact, when they are organized within different political jurisdictions, they are less likely to engage in wars. It is more similar populations, sharing closer preferences over rival goods and resources, that are more likely to fight with each other across national borders.

Interestingly, the historical record also points to significant interactions between conflict within and across borders. For example, two rulers are more likely to fight over a territory (a rival good) if it is inhabited by people who are more similar to those they already rule, as it would later be easier to provide common public goods and policies. In general, a promising direction for future research is the study of the connections and interactions between measures of cultural and political heterogeneity, domestic and international conflict, and the formation and breakup of countries and other political jurisdictions.

References

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003), "Fractionalization", *Journal of Economic Growth*, 8: 55-194.

Arbatli, C. E., Q. Ashraf and O. Galor (2015), "The Nature of Civil Conflict," *working paper*, Brown University.

Caselli, F., M. Morelli and D. Rohner (2014), "The Geography of Inter-State Resource Wars," *Quarterly Journal of Economics*, 130(1), pp. 267-315.

Cavalli-Sforza, L. L., P. Menozzi and A. Piazza (1994), *The History and Geography of Human Genes*, Princeton: Princeton University Press.

Desmet, K., I. Ortuño-Ortín and R. Wacziarg (2012), "The Political Economy of Ethnolinguistic Cleavages," *Journal of Development Economics*, 97 (1): 322–32.

Desmet, K., I. Ortuño-Ortín and R. Wacziarg (2016), Culture, Ethnicity and Diversity, *working paper*, UCLA.

Deutsch, K. W. (1964) "Communication Theory and Political Integration." In *The Integration of Political Communities*, edited by Philip E. Jacob and James V. Toscano. Philadelphia: J. B. Lippincott,

Dyen, I., J. B. Kruskal and P. Black (1992), "An Indo-European Classification: A Lexicostatistical Experiment," *Transactions of the American Philosophical Society*, 82: 1-132.

Esteban, J. and D. Ray (2011), "Linking Conflict to Inequality and Polarization," *American Economic Review*, 101 (4): 1345-74.

Esteban, J., L. Mayoral and D. Ray (2012), "Ethnicity and Conflict: An Empirical Study," *American Economic Review*, 102(4): 1310-1342.

Faten G., G. Palmer, and S. A. Bremer (2004), "The MID3 Data Set, 1993-2001: Procedures, Coding Rules, and Description", *Conflict Management and Peace Science,* 21: 133-154.

Fearon, J. (2003), "Ethnic and Cultural Diversity by Country", *Journal of Economic Growth,* 8: 195-222.

Fearon, J. and D. Laitin (2003), "Ethnicity, Insurgency, and Civil War", *American Political Science Review,* 97(1): 75-90.

Garfinkel, M. R. and S. Skaperdas (2007), "Economics of Conflict: An Overview", in K. Hartley and T. Sandler, eds., *Handbook of Defense Economics*, Amsterdam: North-Holland, 2007, vol. 2, pp. 649–709.

Haas, E. B. (1958) *The Uniting of Europe: Political, Social, and Economic Forces, 1950–1957*. London: Stevens.

Haas, E. B. (1964) *Beyond the Nation State: Functionalism and International Organization*. Stanford: Stanford University Press.

Huntington, S. (1993), "The Clash of Civilizations", *Foreign Affairs* 72(3): 22-49.

Huntington, S. (1996), The Clash of Civilizations and the Remaking of World Order, Simon & Schuster.

Jones, D. M., S. A. Bremer and J. D. Singer (1996), "Militarized Interstate Disputes, 1816–1992: Rationale, Coding Rules, and Empirical Patterns", *Conflict Management and Peace Science,* 15: 163-213.

Mecham, R. Q., J. Fearon and D. Laitin (2006), "Religious Classification and Data on Shares of Major World Religions", *unpublished*, Stanford University.

Montalvo, J. G. and M. Reynal-Querol (2005), "Ethnic Polarization, Potential Conflict and Civil War", *American Economic Review*, 95(3): 796-816.

Schelling, T. (1960), *The Strategy of Con‡ict*, Cambridge, MA: Harvard University Press.

Spolaore, E. (2013), "What is European Integration Really About? A Political Guide for Economists," *Journal of Economic Perspectives*, 27(3), pp. 125-144.

Spolaore, E. (2015), "The Political Economy of European Integration," in *The Handbook of the Economics of European Integration*, Harald Badinger and Volker Nitsc, eds.

Spolaore E. and R. Wacziarg (2009), "The Diffusion of Development", *Quarterly Journal of Economics*, 124(2): 469-529.

Spolaore E. and R. Wacziarg (2013), "How Deep are the Roots of Economic Development?", *Journal of Economic Literature,* 51(2), pp. 325-369.

Spolaore E. and R. Wacziarg (2016a), "Ancestry, Language and Culture", Chapter 6 in Victor Ginsburgh and Shlomo Weber (eds.), *The Palgrave Handbook of Economics and Language*, London:

Palgrave Macmillan.

Spolaore E. and R. Wacziarg (2016b), "War and Relatedness," *Review of Economics and Statistics*, 98(5).

Sumner, W. G. (1906), *Folkways*, New York: Ginn.

Tir, J. (2003), "Averting Armed International Con‡icts Through State-to-State Territorial Transfers", *Journal of Politics*, 65: 1235-1257.

Tir, J., P. Schafer, P. F. Diehl and G. Goertz (1998), "Territorial Changes, 1816-1996: Procedures and Data," *Conflict Management and Peace Science* 16, 89-97.

Weede, E. (1973), "Nation-Environment Relations as Determinants of Hostilities Among Nations," *Peace Science Society (International) Papers*, 20: 67-90.

Appendix: Conflict over Rival Goods with Peaceful Bargaining

In our basic model, the two states engage in conflict when both strongly care about the prize. However, conflict is costly, and both states would be better off if they could agree on an allocation of the prize that replicates the expected allocation from conflict, without bearing the actual costs from violent confrontation. For instance, if the prize is divisible, the two states would be better off by sharing it in proportion to their relative power - i.e., state i would obtain a share equal to π_i and state j would obtain a share equal to $(1 - \pi_i)$. If the prize is indivisible, the states could in principle agree to a lottery where each has a probability of winning the prize equal to its probability of winning the war, therefore saving the costs of going to war. However, even abstracting from issues of imperfect information, it might be extremely difficult to implement such a solution *ex post* (the loser may prefer to go to war after all). Even in the case of a divisible prize, states may have an incentive to unilaterally renege from the bargaining solution, and a war may occur as an equilibrium because each state would be better off fighting than surrendering when the other state fights. In fact, war may be the only equilibrium if each state faces a positive incentive to go to war unilaterally when the other state has agreed to a peaceful negotiation. In the absence of incentives to deviate unilaterally from peaceful bargaining, multiple equilibria may occur: war and peaceful bargaining.

In the latter case, more closely related populations, which share more recent common ancestors and hence may be more similar culturally, linguistically, etc., might be more successful at communicating and coordinating on the efficient equilibrium. If the probability of solving the conflict via peaceful bargaining is indeed higher for more closely related populations, this coordination effect could reduce or offset the main effect stemming from similarity in preferences. Then, the net effect of genetic distance on conflict would be ambiguous. However, coordinating on peaceful bargaining in an anarchic international environment, in the absence of credible commitment technologies, might be relatively rare. Moreover, the hypothesis that more closely related populations are better at coordination is purely speculative, and one could conceive of reasons why coordination may be harder among people who care more strongly about the same rival and excludable goods. Therefore, it is not clear, *ex ante*, whether such coordination effect would reduce or eliminate the main effect of relatedness highlighted in our model. As we have seen in Section 4, the empirical evidence is consistent with the main effect in our model dominating any countervailing effect from coordination on peaceful bargaining. These ideas are formalized below with a simple extension of our basic model of conflict over rival goods.

Consider an extension of the basic model, where peaceful bargaining can follow the choice of actions $\{C, F\}$ - which are now re-interpreted as {challenge, respond to challenge} rather than {challenge, fight}. Assume that, if state j challenges and state i responds to the challenge, each player can choose whether to bargain (B) or to go to war (W). If both choose "bargain," the prize is divided peacefully between the two states, and the two states obtain benefits equal to $\pi_i b_i$ and $(1-\pi_j)b_j$, respectively. That is because we assume that a state's bargaining power depends on its strength should negotiations break down (peaceful bargaining takes place "under the shadow of war").[18] If both states choose W, war follows, with the same payoffs as in the basic model. If state i chooses W while state j chooses B, war also follows, but with the following payoffs:

$$U_i\{W, B\} = (1+\xi)\pi_i b_i - c_i \tag{45}$$

and

$$U_j\{W, B\} = [1-(1+\xi)\pi_i]b_j - c_j \tag{46}$$

where

$$0 < \xi \leq \frac{1}{\pi_i} - 1 \tag{47}$$

The parameter ξ captures the increased probability of winning that results from being the initiator

[18] This is a common assumption in the literature. For example, see Alesina and Spolaore (2005).

of the conflict, in the tradition of Schelling (1960).[19] By the same token, if state i chooses B in the sub-game, but state j chooses W, the payoffs are

$$U_i\{B,W\} = [1-(1+\xi)(1-\pi_i)]b_i - c_i \tag{48}$$

and

$$U_j\{B,W\} = (1+\xi)(1-\pi_i)b_j - c_j \tag{49}$$

Under these assumptions, if one state plays W, the other state is better off playing W rather than B, which implies that $\{W,W\}$ is a Nash equilibrium of the sub-game for all values of the parameters. However, $\{W,W\}$ may or may not be the unique Nash equilibrium. If $\{W,W\}$ is the unique Nash equilibrium, the implications of this extension are the same as the basic model's. If $\{B,B\}$ is also a Nash equilibrium, war may be avoided if both states coordinate on the peaceful equilibrium. Therefore, our model is consistent with Fearon's (1995) discussion of war as emerging from an inability to commit to a Pareto-superior outcome. In our framework both states would be better off if each could commit to play B, but they can do that credibly only if $\{B,B\}$ is also a Nash equilibrium. For the symmetric case ($\pi_i = \frac{1}{2}$ and $c_i = c_j = c$), a necessary and sufficient condition for $\{B,B\}$ to be an equilibrium of the sub-game is

$$\xi \leq \frac{2c}{\min\{b_i, b_j\}} \tag{50}$$

The intuition for the above condition is straightforward: the parameter capturing the unilateral incentives to deviate from bargaining must be small enough for $\{B,B\}$ to be a Nash equilibrium of the sub-game. If $\{B,B\}$ is a Nash equilibrium of the sub-game, it is the unique coalition-proof Nash equilibrium. Three cases are possible: (i) states never coordinate on such an equilibrium even when the condition holds, (ii) states always coordinate on such equilibrium when available, and (iii) sometimes states coordinate, while other times they don't (coordination failure). Cases (i) and (ii) do not modify the implications of the basic model regarding the effect of relatedness on conflict. Relatedness is positively associated with the probability of war whenever conflict occurs - i.e., for all values of ξ in case (i), and for $\xi > \frac{2c}{\min\{b_i,b_j\}}$ in case (ii).

The effect of relatedness on conflict could in principle be modified in case (iii), *if* the likelihood of observing a coordination failure happened to depend on relatedness. For instance, coordination

[19] Analogous results could be obtained by also assuming that the initiator of the conflict faces lower war costs. We abstract from this possibility to keep notation simple.

failure could be more likely across populations that are genealogically more distant, because their norms, habits, languages etc. would tend to be more different, and they might therefore find communication and coordination more difficult. If that were the case, such "coordination failure effect" would reduce the negative correlation between genetic distance and probability of conflict. However, *a priori*, and in the absence of a compelling theory of "equilibrium selection," there is no strong reason to expect that coordination failure would be less likely among more closely related populations. The relationship might even go in the opposite direction: coordination failure could be more likely between more closely related populations - for example, because of mistrust and animosity due to a history of previous conflicts over other rival goods. In the latter case, the effect of relatedness on conflict would be strengthened. As we have seen in the empirical section, the net effect of genetic distance on conflict is negative. This is consistent with two possibilities: (a) coordination failure is not less likely among more closely related populations, (b) coordination failure is less likely among more closely related populations, but this effect is not large enough empirically to offset the main effect of relatedness on conflict highlighted by the basic model.

Table 1 – Effect of Genetic Distance on International Conflict - Probit or IV Probit Estimator
Dependent Variable: dummy for whether a country pair was ever involved in a conflict or war between 1816 and 2001

	(1)	(2)	(3)	(4)	(5)	(6)
	Conflict, univariate specification	**Conflict, baseline specification**	**Conflict, baseline specification IV**	**Conflict, non-contiguous pairs only**	**War, baseline specification**	**War, baseline specification IV**
Genetic distance (GD), weighted	-57.3760** (-17.800)	-19.8786** (-9.317)	-30.6802** (-8.843)	-18.5357** (-9.379)	-6.3389** (-7.478)	-8.6043** (-5.746)
Log geodesic distance		-1.6281** (-5.567)	-1.0182** (-3.090)	-1.4809** (-5.065)	-0.2929* (-2.505)	-0.1728 (-1.349)
Log absolute difference in longitudes		0.1424 (0.731)	-0.0677 (-0.336)	0.1629 (0.842)	-0.0197 (-0.254)	-0.0629 (-0.787)
Log absolute difference in latitudes		-0.1130 (-0.887)	-0.1312 (-1.002)	-0.0729 (-0.614)	-0.1314** (-2.612)	-0.1366** (-2.660)
1 for contiguity		15.4610** (10.095)	16.2256** (5.465)	-	0.8262** (2.701)	0.9060 (1.856)
Number of landlocked countries in the pair		-2.6247** (-9.471)	-2.6311** (-9.566)	-2.4127** (-8.927)	-0.6406** (-5.531)	-0.6500** (-5.635)
Number of island countries in the pair		0.8212** (2.923)	0.8762** (3.005)	0.6967** (2.755)	0.4118** (3.828)	0.4439** (3.711)
1 if pair shares at least one sea or ocean		1.9440** (4.909)	1.9435** (3.799)	1.9330** (5.181)	-0.0154 (-0.128)	-0.0199 (-0.161)
Log product of land areas in square km		0.8940** (18.992)	0.9045** (17.145)	0.7960** (18.528)	0.3132** (17.452)	0.3201** (9.755)
1 for pairs ever in colonial relationship		7.3215** (5.094)	7.6147** (3.175)	8.6303** (6.004)	0.9013* (2.099)	0.9754 (1.463)
1 if countries were or are the same country		1.9512 (1.846)	2.2217 (1.541)	1.6352 (1.229)	1.0952* (2.424)	1.1373 (1.564)
# of observations	13,175	13,175	13,175	12,928	13,175	13,175
Pseudo-R^2	0.075	0.275	-	0.202	0.236	-
Standardized effect (%)	-68.81	-23.84	-36.79	-27.34	-20.57	-27.92

Robust t statistics in parentheses; * significant at 5%; ** significant at 1%. The standardized magnitude refers to the effect of a one-standard deviation increase in genetic distance as a percentage of the mean probability of conflict/war for the sample used in each regression. All coefficients are multiplied by 100.
Source: Spolaore and Wacziarg (2016b), Table 3.

Table 2 – Adding the Effects of Linguistic Distance (LD) and Religious Distance (RD)
Dependent Variable: Dichotomous Indicator of Conflict. Estimator: Probit.

	(1)	(2)	(3)	(4)
	Baseline specification	**Add linguistic distance**	**Add religious distance**	**Add religious and linguistic distances**
Genetic distance (GD), weighted	-29.3281 (8.872)**	-29.1266 (8.792)**	-27.1691 (8.369)**	-27.4118 (8.484)**
Log geodesic distance	-2.4924 (5.374)**	-2.4971 (5.379)**	-2.4498 (5.315)**	-2.4268 (5.291)**
1 for contiguity	22.5037 (10.375)**	22.3377 (10.308)**	21.4007 (10.161)**	21.7116 (10.155)**
Linguistic Distance (LD), weighted	-	-0.8099 (0.659)	-	2.3819 (1.778)
Religious Distance (RD), weighted	-	-	-5.1999 (5.013)**	-5.9958 (5.281)**
Pseudo-R^2	0.250	0.250	0.255	0.255
Standardized effect (%)	-28.050	-27.857	-25.985	-26.217

Robust t statistics in parentheses; * significant at 5%; ** significant at 1%.
The standardized magnitude is the effect of a one standard deviation increase in genetic distance as a percentage of the mean probability of conflict. All coefficients are multiplied by 100 for readability. 10,021 observations used in all columns.
Controls: In addition to reported coefficients, all regressions include controls for log absolute difference in longitudes, log absolute difference in latitudes, number of landlocked countries in the pair, number of island countries in the pair, dummy for pair shares at least one sea or ocean, log product of land areas in square km, dummy for pairs ever in colonial relationship, dummy for countries were or are the same country.
Source: Spolaore and Wacziarg (2016b), Table 5.

Table 3: Fighting over Rival Goods (Oil, Temperate Climate, and Fertile Soil)
Dependent variable: dummy for whether a country pair was ever in conflict in the period specified in row 3

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline (oil sample)	**Oil**	**Baseline (temp. climate sample)**	**Temperate climate**	**Baseline (fertile soil sample)**	**Fertile Soil**
	1945-2001	**1945-2001**	**1816-1900**	**1816-1900**	**1816-1900**	**1816-1900**
Genetic distance weighted	-11.8279** (-6.933)	-7.1885** (-3.184)	-2.0078** (-5.470)	-0.1500 (-0.549)	-1.0450** (-5.396)	-0.0079 (-0.054)
Log geodesic distance	-1.0813** (-5.185)	-1.1553** (-5.361)	-0.0853 (-1.455)	0.0136 (0.331)	-0.0750* (-2.571)	-0.0339 (-1.697)
Interaction of oil producer dummy and gen. dist.		-9.6647** (-3.124)				
Dummy for at least one country in the pair being a major oil producer		1.3988** (3.833)				
Interaction of temperate climate and gen. dist.				-1.2588** (-2.886)		
Dummy for one or more country in the pair with >60% land in temperate zone				0.6627** (8.386)		
Interaction of fertile soil dummy and gen. dist.						-0.8456** (-4.069)
Dummy for one or more country in the pair with >40% fertile soil						0.1381** (7.510)
# of observations	13,175	13,175	10,216	10,216	13,033	13,033
Pseudo-R^2	0.280	0.284	0.261	0.313	0.289	0.322
Standardized effect [a]	-19.690	-28.050	-15.470	-10.850	-8.632	-7.050

Robust z-statistics in parentheses; ** p<0.01, * p<0.05. The standardized magnitude refers to the effect of a one-standard deviation increase in genetic distance as a percentage of the mean probability of conflict/war for the sample used in each regression.
[a]: with interaction effects the standardized magnitude is the *total* standardized effect of genetic distance when the endowment dummy equalso 1.
All coefficients are multiplied by 100 for readability.
Controls: Additional included controls (estimates not reported): log absolute difference in longitudes, log absolute difference in latitudes, dummy=1 for contiguity, number of landlocked countries in the pair, number of island countries in the pair, dummy=1 if pair shares at least one sea or ocean, log product of land areas in square km, dummy=1 for pairs ever in colonial relationship, dummy=1 if countries were or are the same country.
Source: Spolaore and Wacziarg (2016b), Table 6.

Figure 1 - Population Tree

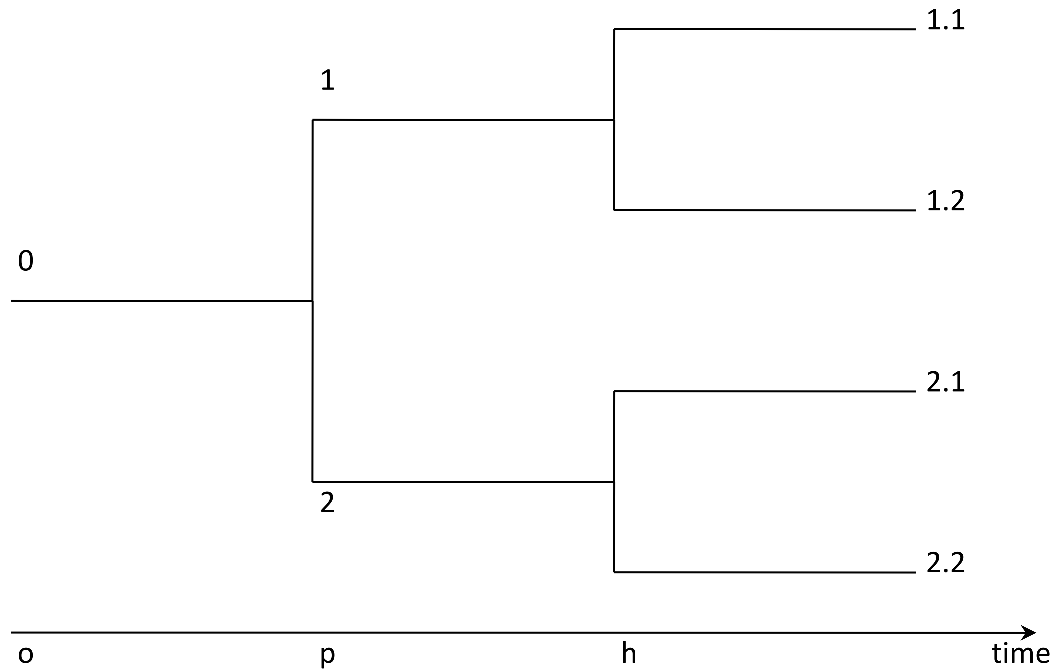
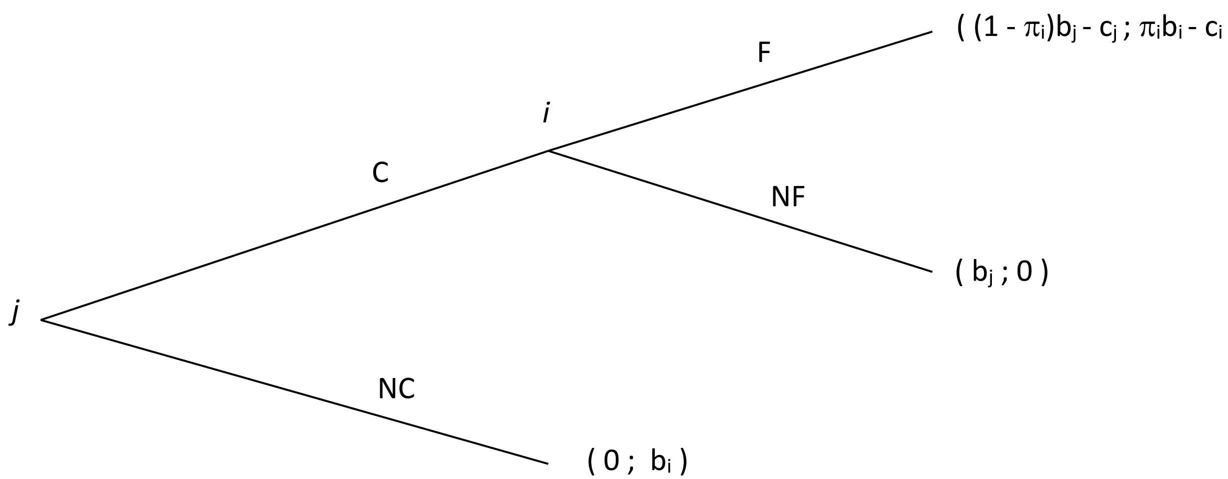


Figure 2 – Extensive-form Game



In parentheses: (state *j*'s payoff ; state *i*'s payoff)

Figure 3 – An Example